STUDY OF ACOUSTIC FEATURES OF NEWBORN CRIES THAT CORRELATE WITH THE CONTEXT

H.E. Baeck[1], M. N. Souza[1,2], *Member*, *IEEE*
[1]Program of Biomedical Engineering, COPPE-Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[2]Eletronic Department of E. E., Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Abstract- **Many researches related to the infant cry analysis intent to estimate the context and/or obtain objective information concerning the physical and emotional condition of newborns. Using several techniques in signal processing, peculiar acoustics features, such as the fundamental frequency and formants, are classically analyzed. However, the findings reveal the existence of some contests with respect to the conclusions. In this article a specific phonologic program was used to analyze the cry signal, aiming to investigate the real significance of some classical frequency domain parameters. The results point out that just two among four studied parameters seem to contribute in the analysis of the cry signal context. Beside the significance of the two parameters in such analysis, the problem complexity indicate that the more researches are necessary to find out new parameters, maybe correlated with psychoacoustic principles.**

***Keywords* - Infants' cry, acoustical features, crying analysis, fundamental frequency.**

I. INTRODUCTION

The cry represents for the human infant one of the few possibilities they have to communicate with their caretakers. Studies have demonstrated that after an experience period, that mothers and nurses acquire an increased ability to recognize the reason of the babies' cry. However, they can not explain the basis of knowledge for such skill and consequently, they are not able to transmit their sensorial experience, concerning to which attributes are important for the performed judgement [1, 2].

Considering that the babies' cry has a communicative intention [3, 4, 5] and that it remains unclear what characteristics of this sound, simply or combined, lead to the perception of the cry's context, it becomes obvious the importance of developing methods to estimate objectively the physiological state of the infant based on its cry, allowing to supply its real needs.

Physiological variables such as, facial expressions, muscular tonus, sleep and suction abilities has been researched as parameters to estimate the needs of the infant [6]. However, during the latest years, the variable "cry" has been used in the major part of the studies [7, 8]. Since 35 years ago, methods that analyze acoustic attributes of the infants' cry have motivated studies that aim to determine its perceptive effects and to search its correlation with the cry context. Studies that correlate the cry with pathologies, using their spectral characteristics as attributes, present clinical applications, able to assist the identification for near missing [9] and early diagnostic of newborns' disturbances and brain malformations [10, 11, 12, 13].

Considering its non-invasive aspect, the acoustic analysis of the cry has been used extensively in studies with newborns and infants with a few days of life. Among the attributes candidates to supply information concerning the infant condition are the duration and the spectral parameters of the cry. The fundamental frequency is generally considered as an outstanding attribute. However its significance and precise role have not been determined yet. According to the authors of one of the most recent publications about cries' analysis [8], there is much contest in published results, begin necessary more studies to clarify the subject.

In the past, technological aspects imposed restrictions to develop efficient analysis techniques [14] to the study of the infants' cry. However in the last few years, with the advent of the informatics' technology producing faster and cheaper computers, important contributions were achieved in understanding the physical and anatomical bases of the cry utterance, as well as its analysis [13,15].

Due to the above considerations concerning the contest in the literature, the present paper shows a systematic study of frequency domain parameters classically stated as important in the literature, i. e., the fundamental frequency F0 and the three first formants F1, F2 and F3, seeking to determinate if these acoustical attributes are really sufficient to categorize the infant cry.

II. METHODOLOGY

In order to provide a better understanding of the contests of this assignment, some terms classically used in the area of crying analysis will be initially defined.

Cry is a term associated to the total duration of the acoustic signal, including all inspirations and expirations, from the beginning of the emission until the infant keep quiet. It is composed by a sequence of *cry units,* as illustrated in Fig. 1, that is defined as the duration of the vocalization only during expiration. When the infant cries we have a sequence of cry units.

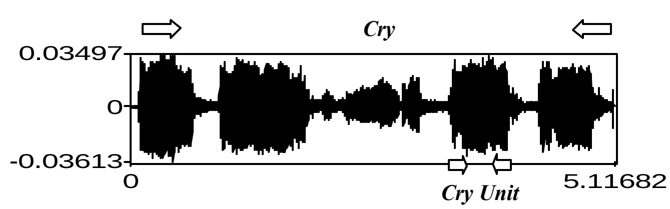

Fig. 1. Indication of the cry and the unit of cry.

This work has been partially supported by Brazilian research agencies CNPq and FAPERJ

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Study of Acoustic Features of Newborn Cries That Correlate With the Context	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Program of Biomedical Engineering COPPE-Federal University Rio de Janeiro, Brazil	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul Turkey. See also ADM001351 for entire conference on cd-rom.

Abstract

Subject Terms

Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU

Number of Pages
4

A. Subjects and Cry Acquisition Protocol

This study has been developed in the Follow Up sector of the Fernandes Figueira Institute, a public pediatric hospital in Rio de Janeiro, were two contexts of cries heave been analyzed: manipulation and pain. The signals called “manipulation cry” were obtained during pediatric evaluation, while the infant was being undressed for weighing. The signals called “pain cry” were obtained during a blood collection for ordinary laboratory tests, at the moment the infant was being punctured in his heel. The puncturing procedure requires not just one, but three consecutive punctures disposed in a triangular way.

The necessary care were taken to control some variables such as: the presence of additional pain stimulus, hunger and loneliness. During the blood collection it is sometimes necessary to compress the punctured area in order to stimulate bleeding, however the nurse was requested not to do so before the end of recording, to avoid the inclusion of a second pain stimulus, other than the pain stimulus being investigated. The infant were breast feed for at least 15 minutes, 20 minutes prior the examination, in order to exclude the possibility of they be hungry during the examination. In order to obtain better record quality, the infants were placed on their back in a canvas bed, maintaining the presence of their caretakers to guarantee the infants would not suffer their absence.

Fifteen infants participated in the pain context and twelve in the manipulation context. All of them born between 25 and 33 weeks of pregnancy.

The cry records were made between when the infant was aged 15 to 30 days. This happened due to the standard procedure in follow up sector and in order to minimize eventual influences that could be caused by differences in the anatomical-physiological nature. As excluding criteria it was considered: infants without the presence of their caretaker, respiratory diseases and that showed excitation before the test.

B. Data Acquisition

The cry signal were acquired using a microphone distant 30,0 cm from the infants’ mouth, being requested to the caretakers not to speak or calm the infant. For the pain cries, the record time was started approximately 5 seconds prior to the puncture with a total duration of 60 seconds. For the manipulation cries, the record time was established the same 60 seconds, starting at the confirmation of the beginning of the cry.

The signals were registered on mini-disc (MD) by means of a digital recorder (Sony model MZ-70) at a sample rate of 44,100 Hz. Despite the digital record of the cry signal, which guarantees the quality of the acquired information, the output of the MD recorder was analog and a second digitalization was necessary . Such digitalization was performed by the software Wave Studio (Creative Labs, version 3.20.0 , 1996), using a Sound Blaster (Creative Labs, model CT 450) with AD/DA resolution of 16 bits, using sample rate of 44,100 Hz.

C. Data Processing

Before the analysis to get the studied parameters, the cry signals were pass-band filtered in the range of 200-5,500 Hz and resampled to 11,025 Hz. The quality of the extracted parameters was guaranteed by the fact the analysis was performed trough a program specially developed to phonological analysis (Praat program, designed by Paulus Petrus Gerardus Boersma from Amsterdam University, Netherlands).

The following parameters were extracted: the fundamental frequency and the tree first formants. The fundamental frequency corresponds to the frequency of the glottal pulse excitation, when the signal is called voiced. Formants are frequency ranges that characteristically contain a concentration of the acoustic energy. The formants represent the natural resonance frequencies of the vocal tract.

For each cry signal the first five cry units were segmented in frames of 25ms and the four parameters (F0, F1, F2 and F3) extracted for each frame. Then, from each parameter value in the frame, a mean value representing the unit of cry was obtained. This procedure was applied in all pain and manipulation context signals.

In order to determine the significance of each extracted parameter, the Student T-Test was realized with the significance level of 0,05. For the parameters whose null hypothesis was not rejected, a correlation test are realized in order to check if the significant parameters correlated each other. The correlation coefficient between each significant parameter and the infant chronological age was also investigated.

III. RESULTS

The tables I and II show the frames’ average results for the studied parameters, for pain and manipulation contexts, respectively. In the end of each raw the mean and the standard deviation for each unit of cry are also showed. For a sake of simplicity each table shows only the results for three cries.

TABLE I
FRAMES’ AVERAGE FROM CRIES OF PAIN CONTEXT

Sig		Unid1	Unid2	Unid3	Unid4	Unid5	ME	SD
Cry 23	F0	655	422	518	416	451	**492.4**	**75.28**
	F1	1670	1958	1968	2100	1965	**1932.2**	**104.8**
	F2	3389	3273	3787	3870	3746	**3613**	**225.6**
	F3	4827	4868	4719	4851	4930	**4839**	**52.8**
Cry 34	F0	397	408	412	472	446	**427**	**25.6**
	F1	1314	1596	1695	1560	1512	**1535.4**	**97.92**
	F2	3131	3504	3640	3329	3330	**3386.8**	**148.1**
	F3	4824	4814	4869	4861	4969	**4867.4**	**41.28**
Cry 120	F0	334	436	448	489	465	**434.4**	**40.16**
	F1	1542	1220	1445	1010	1292	**1301.8**	**153.3**
	F2	3389	3170	3040	3037	3074	**3142**	**110**
	F3	4914	4862	4845	4904	4863	**4877.6**	**25.12**

F0: fundamental frequency (Hz); F1: first formant (Hz); F2: second formant (Hz); F3: third formant (Hz).

TABLE II
FRAMES' AVERAGE FROM CRIES OF MANIPULATION CONTEXT

Sig		Unid1	Unid2	Unid3	Unid4	Unid5	ME	SD
	F0	405	416	405	416	416	**411.6**	**5.28**
Cry	F1	1678	1836	1668	1692	1720	**1718.8**	**47.36**
126	F2	2807	2916	2758	2808	2824	**2822.6**	**37.92**
	F3	4563	4563	4519	4633	4666	**4588.8**	**48.56**
	F0	369	354	328	348	353	**350.4**	**9.92**
Cry	F1	855	807	1001	860	850	**874.6**	**50.56**
136	F2	3434	3468	3486	3446	3470	**3460.8**	**16.64**
	F3	4952	4962	4968	4972	4973	**4965.4**	**6.72**
	F0	450	446	398	492	460	**449.2**	**21.76**
Cry	F1	1206	1235	1381	1231	1169	**1244.4**	**54.64**
137	F2	3308	3294	3328	3303	3296	**3305.8**	**9.76**
	F3	4918	4889	4917	4907	4889	**4904**	**12**

F0: fundamental frequency (Hz); F1: first formant (Hz); F2: second formant (Hz); F3: third formant (Hz).

The tables III and IV show the parameters' mean values to the 14 pain cries and the 12 manipulation cries, respectively. In the bottom of each table the mean and standard deviation, for each parameter, in the two contexts, are also shown.

TABLE III
MEAN AND STANDART DESVIATION FROM THE PARAMETERS EXTRACTED FROM CRY OF PAIN CONTEXT

Sig	F0	F1	F2	F3
Cry 21	486.58	1854	3090	4778
Cry 23	492.76	1932	3613	4839
Cry 28	412.26	1820	3781	4809
Cry 29	402.7	1635	3489	4819
Cry 32	482.54	1822	3234	4816
Cry 34	427.44	1535	3387	4867
Cry 35	545.49	1860	2966	4806
Cry 36	461.78	1530	3551	4715
Cry 114	443.38	1417	3087	4930
Cry 115	428.9	1284	3072	4913
Cry 116	530.33	1272	3187	4905
Cry 118	405.37	1524	3085	4895
Cry 119	462.25	1905	3882	6557
Cry 120	434.78	1302	3142	4878
ME	458.32	1620.85	3326.14	4966.21
SD	44.719	243.81	291.69	461.64

F0: fundamental frequency (Hz); F1: first formant (Hz); F2: second formant (Hz); F3: third formant (Hz).

TABLE IV
MEAN AND STANDART DESVIATION FROM THE PARAMETERS EXTRACTED FROM CRY OF MANIPULATION CONTEXT

Sig	F0	F1	F2	F3
Cry 126	412.13	1719	2823	4589
Cry 127	432.29	1873	3209	4672
Cry 136	351	875	3461	4965
Cry 137	449.65	1245	3306	4904
Cry 139	459.06	1459	3197	4910
Cry 140	367.89	1370	3266	4926
Cry 141	461.04	1264	3242	4896
Cry 146	441.39	1109	3243	4967
Cry 147	367.87	1067	3299	4952
Cry 148	434.56	1658	3290	4900
Cry 148B	427.9	1678	3389	4913
Cry 149	370.81	1297	3512	4973
ME	414.63	1384.5	3269.75	4880.58
SD	39.65	300.78	171.12	121.32

F0: fundamental frequency (Hz); F1: first formant (Hz); F2: second formant (Hz); F3: third formant (Hz).

The table V shows the results of the Student T-test, were Ho=1 indicates that the null hypothesis was rejected and Ho=0 indicates it was not rejected, for a significance level of 0,05.

TABLE V
STUDENT T-TEST

Features	Ho	Significance
F0	1	0.015
F1	1	0.036
F2	0	0.562
F3	0	0.539

F0: fundamental frequency (Hz); F1: first formant (Hz); F2: second formant (Hz); F3: third formant (Hz).

The table VI shows the result of the correlation tests between the significant parameters and also between each parameter and the infant chronological age.

TABLE VI
CORRELATION COEFICIENT

Parameters	Context	Pearson Coef.
	D	0.26
F0 e F1	M	0.41
	D	0.43
F0 e id	M	0.21
	D	0.32
F1 e id	M	-0.10

F0: fundamental frequency (Hz); F1: first formant (Hz); id: idade em meses; D: dor; M: manipulação.

IV. DISCUSSION AND CONCLUSION

The comparison between the mean and standard deviation shown in tables III and IV point out the parameters F2 and F3 do not play a significant role in the discrimination of the pain and manipulation contexts. The same can be seen by the results of the Student T-test, shown in Table V, were the null hypothesis was not rejected for either the two parameters.

The fundamental frequency F0 has been described in the literature as the principal feature that affects the auditory perception of the babies' cry. In the present study the importance of such parameter was confirmed. Tables III and IV show a significance change in the average and standard deviation to the F0 parameter, as well as to the F1 parameter, pointing out both of them contribute to the context identification. The results of the Student T-test in Table V agree with such conclusion.

Since the results indicated the F0 and F1 were significant parameters a correlation test was performed to verify if two parameters have really been detected, or if just one of them was significant and the other linear correlated with the first. Table VI shows that F0 and F1 are not correlated each other and both of them seem to be independent parameters that can be used to identify the cry context. The parameter F1 has been reported as correlated to the lip area [16] and consequently it is reasonable to considered the pain context exhibits a greater value of F1, when compared with manipulation context, because the baby normally opens more its mouth in the first case. The increase in the F0 parameter from the manipulation to the pain context can be explained by the fact an increasing in vocal folds strength, due to the change in the emotional state, leads to a higher vibration frequency.

This paper investigated features connected to the frequency domain pointing out F0 and F1 are important parameters to cry analysis systems. Consider that this two parameters alone are sufficient to develop a total reliable system seems not to be reasonable due the complexity of the problem. It is necessary to continue the investigations for new features, maybe relate the temporal domain and/or psychoacoustical effects.

ACKNOWLEDGMENT

The authors would like to tank professor Paulus Petrus Gerardus Boersma from the Amsterdam University by the permission to use his program Praat. They also wish to thank Dr. C. A. M. Sá, Dr. M. D. B. Méio and R. S. Santos from the Follow Up sector of the Fernandes Figueira Institute for their cooperation with the experimental setup.

REFERENCES

[1] M. D. S. Ainsworth, "Object relations, dependency, and attachment: a theoretical review of the infant-mother relationship," *Child Development*, vol. 40, pp. 969-1025, 1969.

[2] J. E. Drummond and M. L. McBride, "The development of mother's understanding of infant crying,". *Clinical Nursing Research*, vol. 2, pp. 396-411, 1993.

[3] M. Papousek, "Determinants of responsiveness to infant vocal expression of emotional state," *Infant Behavior Development*, vol. 12, pp. 507-524, 1989.

[4] A. Protopapas and P. D. Eimas, "Perceptual differences in infant cries revealed by modifications acoustic features," *J. Acoust. Soc. Am.*, vol. 102, No 6, pp. 3723-3734, Dec 1997.

[5] D. L. Boero, A. M. C. Volpe, C. Bianchi and C. Lenti, "Newborns Crying in different contexts: discreate or graded signals ?," *Perceptual and Motor Skills*, vol. 86, pp. 1123-1140, 1998.

[6] Y. Skogsdal, M. Eriksson and J. Schollin, "Analgesia in Newborns given oral glucose" *Acta Paediatrics*, vol. 86, pp. 217-220, 1997.

[7] J. Lawrence, D. Alcock, P. Mcgrath,, J. Kay, B. MacMurray and C. Dulberg, "The development of a tool to assess neonatal pain," *Neonatal Network*, vol. 6, pp. 59-65, 1993.

[8] P. Runefors, E. Arnbjornsson, G. Elander and K. Michelsson, "Newborn infants'cry after heel-prick: analysis with sound spectrogram," Acta Paediatr, vol 89, No 1, pp. 68-72, 2000.

[9] B.M. Lester and C. F. Boukydis, "No language but a cry". In: H. Papuosek, U. Jurgens, M. Papuosek (eds.), "Nonverbal vocal communication: comparative and developmental approaches", Cambridge: University Press, pp. 145-173, 1992.

[10] R. D. Kent and A. D. Murray, "Acoustic features of infant vocalic utterances at 3, 6, 9 months," *J. Acoust. Soc. Am.*; vol. 72, pp. 353-365, 1982**.**

[11] L.Gray, "Signal Detection analysis of delays in neonates vocalizations,". *J. Acoust. Soc. Am.,* vol. 82, pp.1608-1611, 1987.

[12] A. Fort, A. Ismaelli, C. Manfredi, and P. Bruscaglioni, "Parametric and non-parametric estimation of speech formants: aplication to infant cry,". *Medical Engineering and Physics*, vol. 18, pp. 677-691, 1996.

[13] K. Michelsson and P. Sirviö, "Cry analysis in congenital hypothyroidism," *Folia Phoniatrica*, vol. 28, pp. 40-47, 1976.

[14] Q. B. Xie, R. K. Ward and C. A. Laszlo, "Automatic assesment of infants' levels-of-distress from the cry signals" *IEEE Trans. On Speech and Audio Proc.*, vol 4,No.4, pp. 253-265, July 1996.

[15] B. F. Fuller and Y. Horii, "Spectral energy distribution in four types of infant vocalizations," *J. Commun. Disord.*, vol 21, pp. 251-261, 1988.

[16] P. P. G. Boersma, Funcional Phonology, Holland Academic graphics, Amsterdam, NL, 1998, pp. 509.